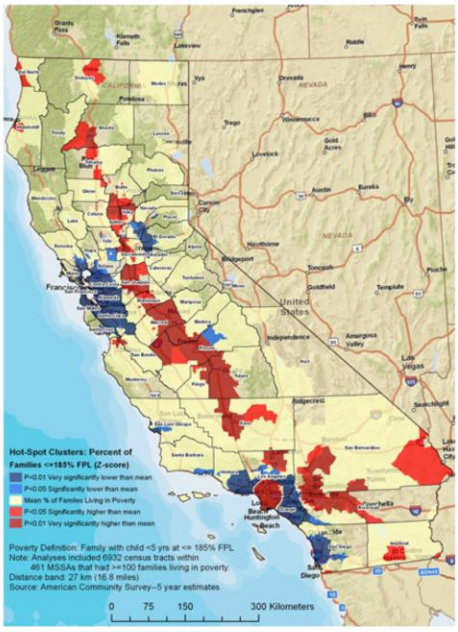
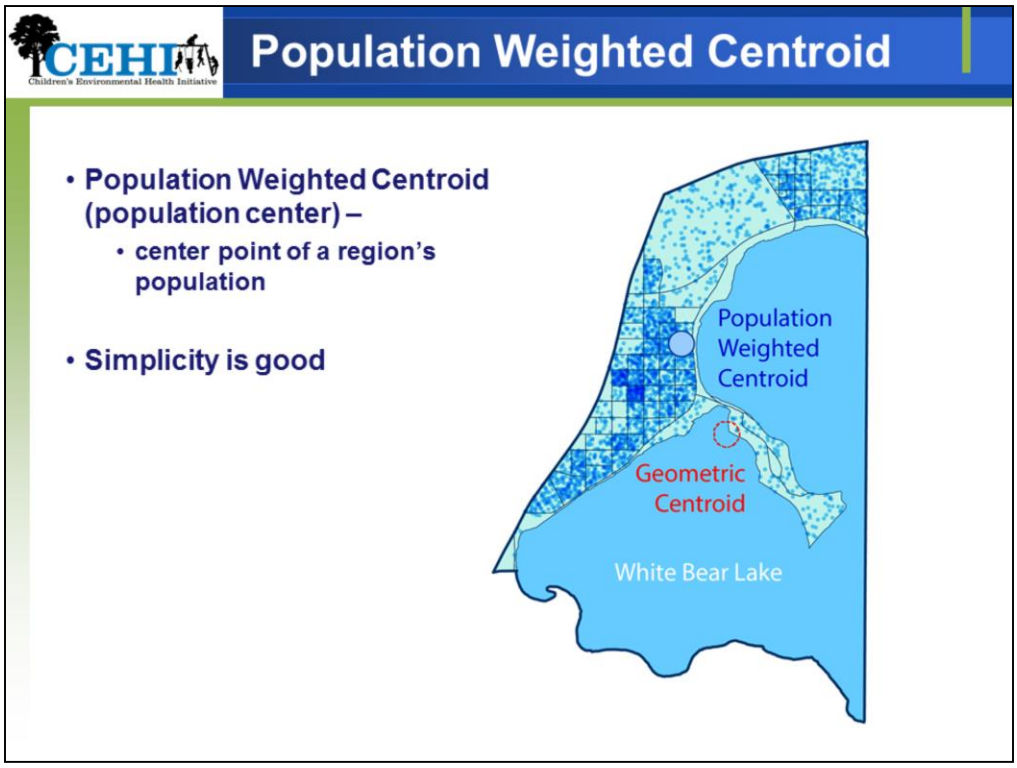
Analyzing patterns in your data

Using GIS Training to Address Blood Pressure Medication Adherence

Learning Objectives

- Conceptualize where your population lives
- Understand scale issue in spatial studies
- Get to know spatial autocorrelation
- Examine and interpret hotspot analysis





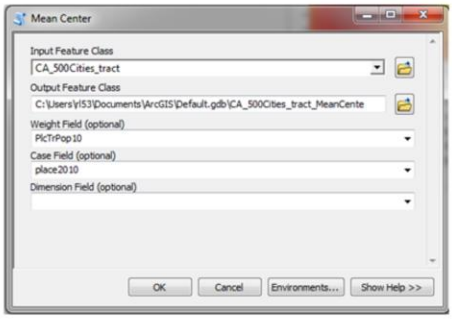
Points are easier to work with. The relationship to other feature can be simplified.

Generate your own Population Weighted Centroid

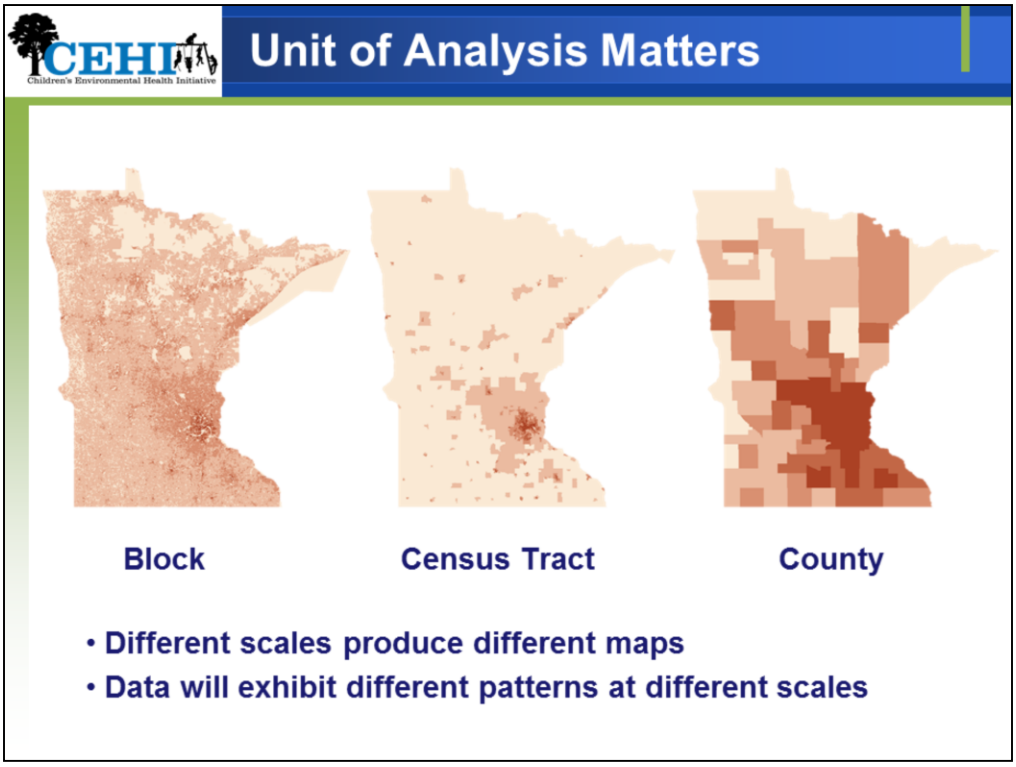
- Use population in sub-unit as weight
- Make sure you have the identifier for the polygons
- Other variables can be used as weight

$$(X, Y) = \frac{\sum_{i=1}^{n} p_i \, (x_i, y_i)}{\sum_{i=1}^{n} p_i}$$

* *where (X, Y) is the location for PWC, (x_i, y_i) is population location in the region, and p_i is the population for point i*



Population weighted centroid (population center) data is available at census website: https://www.census.gov/geo/reference/centersofpop.html



Maps showing here are population density for MN at different scale. Maps are displaying in quintile classification across all maps.

Spatial Autocorrelation

- Spatial data is rarely independent
- Tobler's First Law of Geography
- Neighborhood should be taken into consideration

Tobler's First law of Geography: Everything is related to everything else, but near things are more related than distant things

Hotspot Analysis

- Use Getis-Ord G* index

$$G_i^* = \frac{\sum_{j=1}^{n} \omega_{ij} x_j}{\sum_{j=1}^{n} x_j}$$

- Compares observed G* with expected G*

$$Z_i^* = \frac{G_i^* - \overline{G_i^*}}{\sqrt{Var(G_i^*)}}$$

- Shows statistical significance

* where x_j is the local measure (prevalence), ω_{ij} is spatial weight matrix.

- Hot spots in red and cold spots in blue

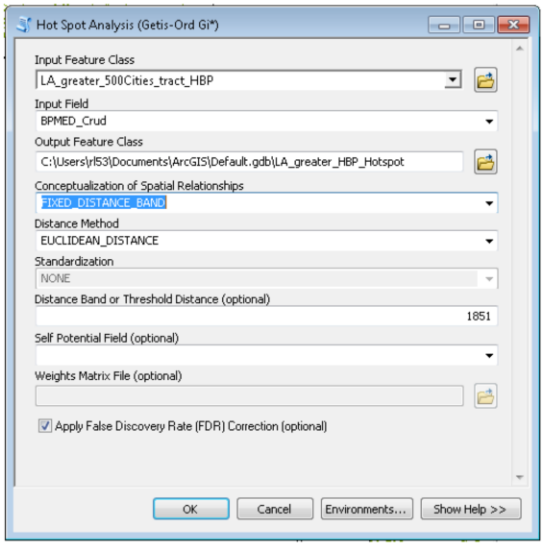
Hotspot analysis uses vectors (not rasters)

Use Getis-Ord G* index (taking neighborhood value into account)

Density can tell you where clusters in your data exist, but not if your clusters are statistically significant

Considerations for Hotspot Analysis

- **Hotspot analysis is a diagnostic tool**
- **Spatial relationship settings is crucial**
  - **Distanced based: Fixed Distance Band, and etc.**
  - **Polygon Contiguity: edge/ edge corner**
  - **Get Spatial Weights From File**
- **Multi-test correction**



Distanced based - Etc. : Inverse Distance, Inverse Distance Squared, Zone of Indifference

Euclidean distance weighted – Closer features are weighed more heavily than features that are further away

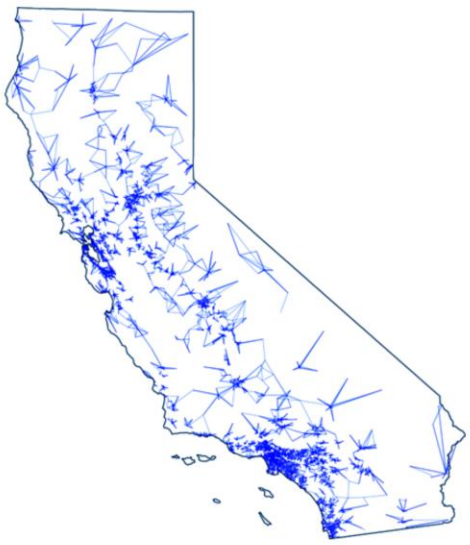
Fixed Distance Band – Every feature within a fixed distance is included, every feature outside that distance is excluded

Zone of Indifference – Combination of Inverse Distance and Fixed Distance Band

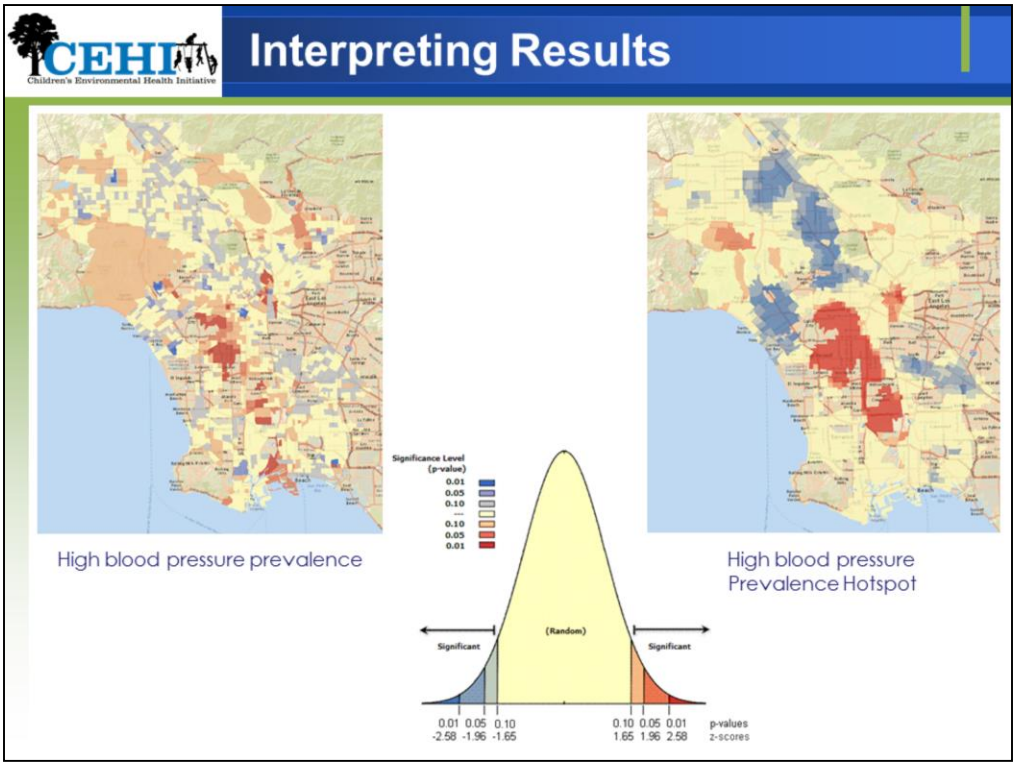
Polygon Contiguity – Only features that share a border are included

Define Study Scale

- Data must be projected!
- Distance unit needs to match the projection
- "Knowledge" is the key
- Useful information is the goal
- Driving distance could be used to estimate community



In the exercise, we want to look at hot spots/ cold spots for high blood pressure prevalence in greater LA region. We used average of the driving distance to nearest 5 pharmacies from census population weighted centroids. Specifically speaking, 3529 meters

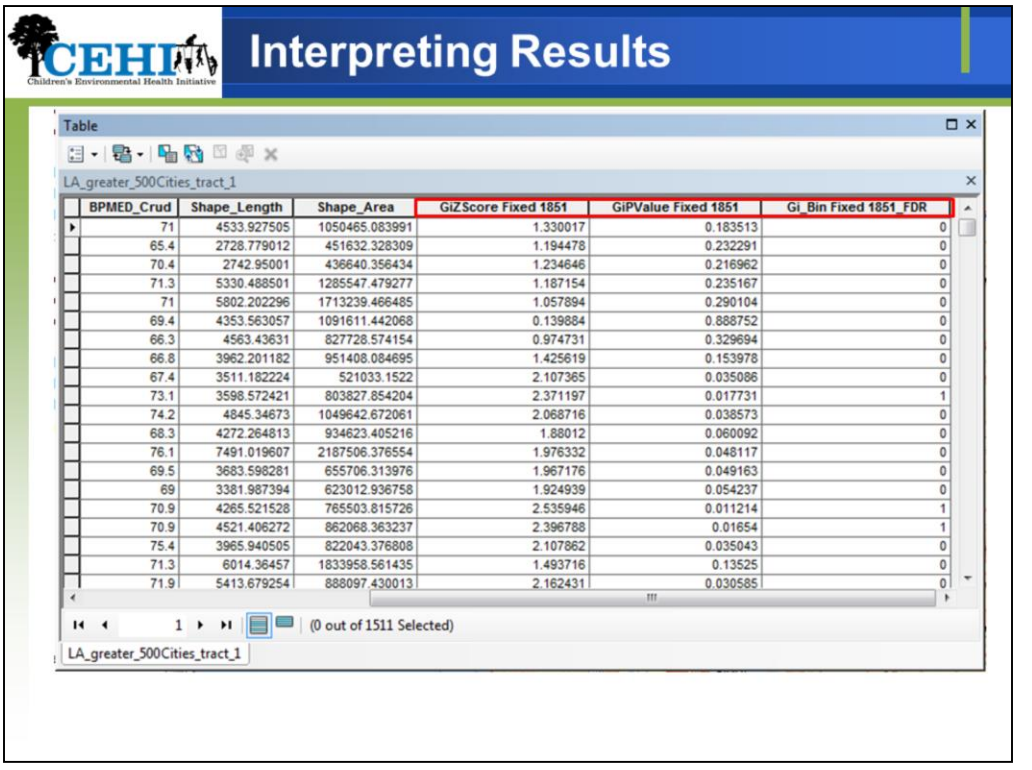


A high Z score and small P value for a feature indicates a significant hot spot.

A low negative Z score and small P value indicates a significant cold spot. The higher (or lower) the Z score, the more intense the clustering.

A Z score near zero means no statistically significant spatial clustering.

Interpreting Results



Resulting z-scores and p values are stored in the attribute table of your output.

The “Bin” attribute field contains the categories you can use to display hotspots (considers the z-scores and p values to create the categories).

You don’t want to compare either Z score and P value, Gi_bin is the place to look at